Large geographic range size reflects a patchwork of divergent lineages in the long-toed salamander (*Ambystoma macrodactylum*)

J. A. LEE-YAW & D. E. IRWIN

Department of Zoology, University of British Columbia, Vancouver, BC, Canada

Abstract

For northern taxa, persistence in multiple vs. single Pleistocene refugia may have been an important determinant of contemporary range size, with larger ranges achieved by species that colonized the north from several glacial refugia. Under this hypothesis, widespread species are expected to demonstrate marked phylogeographic structure in previously glaciated regions. We use a genome-wide survey to characterize genetic structure and evaluate this hypothesis in the most widely distributed salamander in the Pacific Northwest, the long-toed salamander (*Ambystoma macrodactylum*). Patterns of variation based on 751 amplified fragment length polymorphism (AFLP) loci and mitochondrial sequence data were concordant and support the recognition of at least four distinct lineages of long-toed salamander. The distributions of these lineages indicate that multiple refugia contributed to the species' large contemporary range. At the same time, with up to 133 AFLP bands differing between lineages and levels of sequence divergence ranging from 2.5 to 5.8%, these lineages would be considered separate species by some definitions. Such splitting would partition the large geographic range of the long-toed salamander into several relatively restricted ranges. Our results thus also underscore the potential for estimates of geographic range size to vary considerably depending on the taxonomic treatment of cryptic lineages.

Introduction

That most species have relatively small geographic distributions is a well-established pattern in macroecology (reviewed by Gaston, 2003). Identifying the factors that have permitted some species to achieve comparably large distributions is a major goal in ecology, with leading hypotheses invoking traits such as dispersal ability (Lester *et al.*, 2007; Leger & Forister, 2009), body size (Cambefort, 1994; Gaston & Blackburn, 1996) and niche breadth (Garcia-Barros & Romo Benito, 2010). Recently, Shafer *et al.* (2010) demonstrated an association between contemporary range size and persistence in multiple vs. single glacial refugia in western North America, adding a potentially important historical dimension to understanding large geographic ranges in

previously glaciated regions. Specifically, this pattern suggests that persistence in multiple glacial refugia and the patching together of the post-glacial ranges of several populations (Hewitt, 2004) was instrumental to the formation of large ranges in northern areas.

Isolation and divergence in multiple glacial refugia during the Pleistocene has also been implicated in speciation in some taxa (Weir & Schluter, 2004; Levsen *et al.*, 2012; but see Barber & Jensen, 2012 and references therein). Thus, the pattern observed by Shafer *et al.* (2010) raises the additional possibility that some widespread species actually consist of genetic groups that, following taxonomic scrutiny, will warrant species status. Corresponding reductions in described geographic ranges may have important implications for comparative studies that rely on estimates of range size to test hypotheses for variation in range size (see Loxdale *et al.*, 2011 for related discussion).

Characterizing genetic structure in northern areas is an important first step towards determining whether widespread species represent patchworks of divergent

Correspondence: Julie A. Lee-Yaw, University of British Columbia, Department of Zoology, 6270 University Blvd., Vancouver, BC, V6T 1Z4, Canada. Tel.: +1 778 846 7391; fax: +1 604 822 2416; e-mail: lee-yaw@zoology.ubc.ca

groups and the extent to which such groups warrant taxonomic recognition. Although many existing phylogeographic studies speak to this goal, most have been conducted using just a handful of genetic markers. The potential for stochasticity and selection to cause patterns for any single marker to deviate from the overall history of closely related groups limits the conclusions that can be drawn from studies using a small number of markers (Irwin, 2002, 2012; Rauch & Bar-Yam, 2004; Dowling *et al.*, 2008). Methods that survey many markers throughout the genome overcome this concern and inherently assess the degree of divergence between taxa. Because the use of a limited number of genetic markers often leads to the underestimation of genetic diversity (e.g. Irwin *et al.*, 2009; Mila *et al.*, 2010; Wiens *et al.*, 2010), revisiting phylogeographic structure with these high-resolution methods may be particularly pertinent for determining whether widespread species represent single cohesive genetic entities or composites of genetically distinct groups.

The long-toed salamander (*Ambystoma macrodactylum* Baird, 1849) exemplifies the case of a widespread species for which phylogeographic patterns and the extent of diversification remain unclear, despite previous attention. Currently considered one species, the long-toed salamander is the most widely distributed salamander in the Pacific Northwest, a region that is both geologically and ecologically complex (Brunsfeld *et al.*, 2001). The potential for the species' wide range to consist of multiple distinct lineages was first suggested by Mittleman (1948) and later Ferguson (1961), whose work examining vomerine tooth count, allometry and strip pattern and coloration in the southern portion of the species' range, led to the recognition of five subspecies of long-toed salamander (Ferguson, 1961; Fig. 1). More recently, two studies employing a limited number of genetic markers (mtDNA: Thompson & Russell, 2005; mtDNA and six nuclear loci: Savage, 2008) have provided some evidence of genetic differences between named subspecies. However, different patterns observed among these studies and among markers within the study by Savage (2008), as well as discrepancies between genetic data and morphological subspecies (Thompson & Russell, 2005; Savage, 2008) make the interpretation of subspecies difficult. Furthermore, limited sampling to date, especially in northern areas, restricts the conclusions that can be drawn about the distribution of distinct lineages and the phylogeographic structure of the species in previously glaciated regions. The extent to which previously described groups are representative of genome-wide patterns of differentiation thus requires further assessment.

In this study, we undertake a comprehensive genetic survey to better understand diversification in the long-toed salamander. Comparing phylogeographic patterns observed using amplified fragment length polymorphism (AFLPs) to those identified by mitochondrial
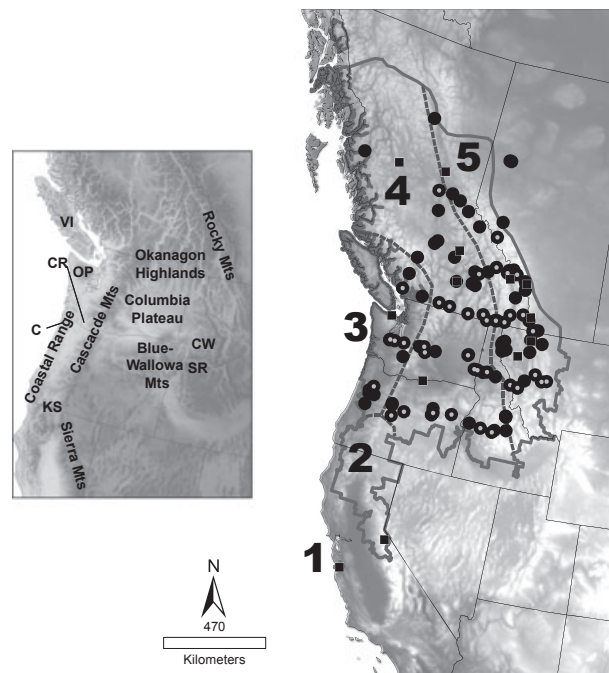

Fig. 1 The geographic distribution of *Amybstoma macrodactylum* (thick grey line). Described subspecies' boundaries are depicted as dashed lines: (1) *A. m. croceum,* (2) *A. m. sigillatum,* (3) *A. m. macrodactylum,* (4) *A. m. columbianum* and (5) *A. m. krausei.* Circles represent sites where both amplified fragment length polymorphism (AFLP) and mtDNA data were assayed (*n* = 1–11 individuals/site) in the present study. Those sites with samples that were also included in a larger AFLP data set (see Materials and Methods; Fig. 3) are indicated with a small grey circle within the larger black circle. Squares represent sites where only mitochondrial DNA (mtDNA) sequence data were collected (*n* = 1–4 individuals/site). Inset: Major geographic features and the locations of previously reported glacial refugia as follows: VI, Vancouver Island; OP, Olympic Peninsula; CR, Columbia River Drainage; C, Coastal Mountains; CW, Clearwater River Drainage; SR, Salmon River Drainage; KS, Klamath-Siskiyou Mountains.

DNA, we ask whether genome-wide patterns of differentiation support the existence of multiple lineages of long-toed salamander. We simultaneously use our genetic survey to determine the distribution of major genetic groups, with a focus on northern parts of the species' range. In doing so, we clarify the extent to which one of the largest geographic ranges in the Pacific Northwest represents a truly remarkable feat of colonization by a single taxonomic entity vs. a patchwork of multiple, divergent lineages.

Materials and methods

Sampling

Tail clips were collected during the spring and summers of 2008 through to 2010. Other researchers and the

Museum of Vertebrate Zoology (University of California, Berkeley) provided additional samples. In total, 403 individuals from 122 sites across the range of the species were sampled (Fig. 1; Table S1).

AFLP data collection

Total genomic DNA was extracted using a standard phenol–chloroform protocol. We generated AFLP profiles for 378 individuals from 108 sites (Fig. 1). The 25 remaining individuals were only included in the mitochondrial data set (below) due to either poor DNA quality or because they were obtained after the ALFP data had been generated. We used the AFLP protocol of Vos *et al.* (1995), modified according to Toews & Irwin (2008), for the digestion, ligation and preamplification steps. Selective amplification followed the protocol of Clarke & Meudt (http://clarkeresearch.org/aflp_2012-01-26/AFLP_Protocol.pdf). To reduce the complexity of banding patterns resulting from the large size of *Ambystoma* genomes, we included two selective amplification steps (Voss & Shaffer, 1997). Using these protocols, we generated two AFLP data sets. Our most inclusive data set of 378 individuals was generated using one selective primer combination in the final selective amplification (Table 1). A subset of 70 of individuals from 39 populations (Fig. 1) was then screened for an additional five selective primer combinations. This second data set thus represents more comprehensive sampling of the genome for a representative set of individuals.

Fragment detection was conducted on an automated ABI 3100. AFLP electropherographs from each primer combination were imported into Peak Scanner v.1.0 (Applied Biosystems 2006, Carlsbad, CA, USA) for initial analysis. Peak sizes were called using default settings except for the application of lite smoothing

recommended for automatic scoring (Arrigo *et al.* 2012). The automatic binning and scoring algorithm implemented in RawGeno v. 2.11.1 was then used to analyze peaks exported from PeakScanner for each primer pair. The bin detection range was 50 to 400 bp to minimize homoplasy associated with very small fragments (Vekemans *et al.*, 2002) and detection issues associated with drop-off at larger fragment sizes. Individuals were scored as having a band 'present' for a bin if a corresponding peak exceeded 80 rfu. To reduce the number of uninformative bins, we eliminated bins for which fewer than five individuals were scored as having the band present. Additionally, 5–13% of individuals (depending on primer pair), replicated from DNA extraction, were included in the set of samples for each primer pair. We set the repeatability filter in RawGeno to remove bins that were < 80% repeatable across these replicates (cut-off recommended by N. Arrigo, pers. com.). After filtering bins, our two AFLP data sets included 177 loci (average repeatability 92%) and 751 loci (average repeatability of 91.3%) respectively.

mtDNA data collection

For comparison with the AFLP data, we sequenced the mtDNA cytochrome b gene from 142 individuals. Primers were designed from conserved regions of the mitochondrial genomes of *A. laterale* (GENBank Accession NC_006330; Mueller *et al.*, 2004), *A. mexicanum* (GENBank Accession NC_005797; Arnason *et al.*, 2004) and *Plethodon petraeus* (GENBank Accession NC_006334; Mueller *et al.*, 2004): AmbPleth_cytb-F 5′ACYGRAAC CYTTGACMTGAA; AmbPleth_cytb-R 5′YCRRTTTTCGRC TTACAAGG. Sequences were obtained using nested primers for two museum specimens from the Sierra Nevada Mountains and Santa Cruz (California) that were of reduced quality and thus not included in the AFLP data set. PCR was carried out in 25-μL reactions consisting of 0.25 μL dNTPs (10 μM), 2.5 μL 10x reaction buffer (Invitrogen), 0.75 μL MgCl (50 mM), 1 μL each primer (10 μM), 18.4 μL ddH$_2$0, 1 μL TAQ (5000 U mL^{-1}: New England Biolabs, Ipswich, MA, USA) and 1 μL DNA (25 ng μL^{-1}). Amplification involved initial denaturation at 94 °C for 2 min, 35 cycles of denaturation at 94 °C for 30 s, annealing at 56 °C for 45 s and extension at 72 °C for 1 min and 10 s and a final extension at 72 °C for 7 min. PCR products were sequenced by the Genome Quebec Innovation Centre at McGill University on an automated ABI 3730XL. The resulting chromatographs were verified by eye. Ten individuals (including the two California samples) were initially sequenced in both directions and compared in BioEdit Version 7.0.5.3 (Hall, 1999). As all sequences were unambiguous and identical in both directions, we sequenced the remaining individuals using the reverse primer only. Use of this primer resulted in 729 bp of sequence data. Sequences were

Table 1 Amplified fragment length polymorphism primer combinations used to generate loci for addressing phylogeographic structure in the long-toed salamander. The first primer combination listed was that used to generate AFLP profiles for all individuals; the remaining five combinations were used to generate additional data for a subset of 70 individuals.

EcoRI primer (*NNN-3′)	MseI primer (†NNN-3′)	Dye‡	Number of polymorphic fragments (after filtering)
AGCA	CATA	NED	177
AGCG	CATC	PET	178
AGCA	CATG	NED	95
AGCG	CATA	PET	105
AGCT	CATA	VIC	101
AGCC	CATC	FAM	95

*EcoRI primer: GACTGCGTACCAATTC
†MseI primer: GATGAGTCCTGAGTAA
‡Applied Biosystems Canada (Toronto, ON, Canada)

manually verified and edited in BioEdit v. 7.0.5.3 (Hall, 1999) and aligned using Clustal (Thompson *et al.*, 1994).

Following our analysis of the sequence data (below), we obtained mtDNA group membership information for the remaining 261 individuals included in the AFLP data set using PCR-RFLP (Appendix S2). Briefly, samples were amplified for a 597-bp section of the cyt b gene described previously and then digested with up to three restriction enzymes that allowed us to unambiguously assign individuals to mitochondrial group based on fixed SNPs observed in the sequence data. Thus, mtDNA data were available for all sites included in the study, allowing us to compare mtDNA lineage boundaries with those observed for the nuclear genome.

Evaluating genetic structure

We used two methods to determine whether genome-wide patterns revealed by the AFLP data support distinct lineages of long-toed salamander. For both AFLP data sets, the Bayesian clustering algorithm implemented in STRUCTURE (Pritchard *et al.*, 2000; Falush *et al.*, 2007) was used to calculate posterior probabilities of the assignment of individuals to 1–12 groups (K) using the admixture model of ancestry. For each K, we conducted 10 runs of 1 000 000 MCMC generations with the first 500 000 generations discarded as burn-in. The optimal value of K was determined by examining the log probability of the data as well as using the method outlined by Evanno *et al.* (2005). For the AFLP data set generated from six selective primer combinations (i.e. the largest survey of the genome), we also examined similarity between individuals using a multi-dimensional scaling analysis based on Jaccard distances calculated using the VEGAN package in R (Oksanen *et al.*, 2011).

Relationships amongst individuals based on mtDNA were explored using the haplotype network estimation procedure implemented in TCS version 1.21 (Clement *et al.*, 2000). Due to the sensitivity of the analysis to ambiguous data, we excluded characters for which any individual had ambiguous base calls. Network estimation was performed on two data partitions. The first partition of 600 unambiguous characters excluded the sole representative from the Sierra Nevada Mountains in California as we had limited sequence data for this individual and we wished to infer relationships between haplotypes with the largest number of characters possible. The second partition was a 489 unambiguous character subset of the data that allowed inclusion of this individual for purposes of determining its likely position in the network.

Assessing degree of divergence

We used the 70 individuals for which we had the largest amount of AFLP data to assess genetic divergence

between the major genetic groups. ARLEQUIN 3.11 (Excoffier *et al.*, 2005) was used to calculate the average distance between groups in terms of the number of bands and to perform an analysis of molecular variance (AMOVA) to evaluate the percentage of AFLP variation attributed to differences within vs. between major groups. These calculations required *a priori* assignment of individuals to groups. To avoid circularity, we assigned individuals to groups based on the major mitochondrial breaks suggested by the haplotype network rather than using the nuclear DNA groups suggested by STRUCTURE. We also calculated average uncorrected and corrected pairwise sequence divergence between the mitochondrial groups using the APE package in R (Paradis *et al.*, 2004). Corrected differences were calculated using the Tamura & Nei (1993) model of sequence evolution with gamma parameter.

Results

Genetic structure

Substantial genetic structure was observed in the AFLP data sets. Examination of mean ln probabilities and calculation of ΔK (Evanno *et al.*, 2005) following the STRUCTURE analyses revealed a clear peak at $K = 5$ for both data sets (Fig. S1). Four of these groups are geographically distinct, partitioning the species into a Coastal-Cascade group, a North-Central group, a Rocky Mountains group and a Central Oregon Highlands group (Fig. 2). The proportion of each individual's ancestry from the fifth group was generally low (although there were a few exceptions in the analysis of 177 loci: Fig. 2). There was no clear geographic structure associated with this fifth group.

The multi-dimensional scaling analysis of 751 AFLP loci resolved three of the genetic groups observed in the STRUCTURE analysis: the Coastal-Cascade, North-Central and Rocky Mountains groups (Fig. 3). Individuals from the Central Oregon Highlands clustered with Coastal-Cascade individuals, although there was slight divergence between these groups along Dimension 1 of the MDS scaling plot. Close association between individuals from the Central Oregon Highlands and Coastal-Cascade Mountains is also suggested by the clustering of these individuals as one group in the STRUCTURE analysis of $k = 4$, one less than the optimal value (plot not shown; see Prunier & Holsinger, 2010 for this approach).

Strong phylogeographic structure was also evident in the mitochondrial genome. Specifically, haplotype network estimation using the data set with the larger number of unambiguous characters resulted in six distinct networks that could not be connected within the limits of statistical parsimony, in this case 10 mutational steps (Fig. 4). Four of these networks coincide with the groups resolved by the STRUCTURE analysis (Coastal-Cascade, North-Central, Rocky Mountain and Central Oregon
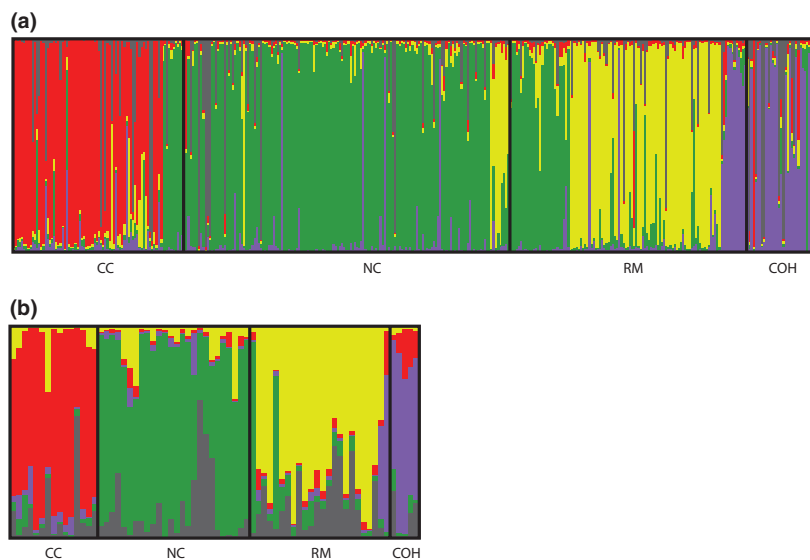
Fig. 2 Population structure of long-toed salamanders detected using the Bayesian assignment method of structure in analysis of 177 AFLP loci for 378 individuals (a) and 751 AFLP loci for a subset of 70 individuals (b). Individuals are represented as bars (see Table S1 for population order for a), with the posterior probability of assignment to each of five clusters (the optimal value of K for both data sets) represented by the different colours. Mitochondrial (mtDNA) group membership based on the haplotype network analysis is indicated below the plot with mtDNA groups separated by thick black lines. CC, Coastal-Cascade; NC, North-Central; RM, Rocky Mountains; COH, Central Oregon Highlands. Individuals are lined up such that those closest to the breaks separating mtDNA groups come from populations closest to the contact zones between lineages (Appendix S1).
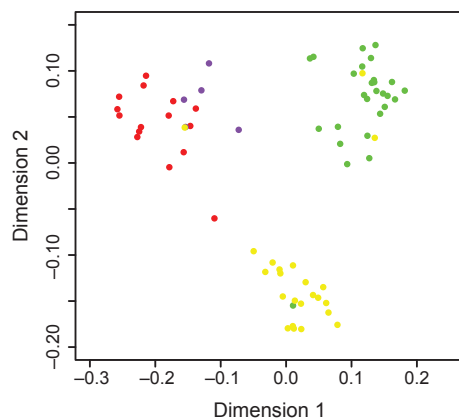

Fig. 3 Multi-dimensional scaling plot of 70 long-toed salamanders based on Jaccard distances calculated from 751 AFLP loci. Individuals are colour-coded according to mitochondrial DNA group membership from the haplotype network analysis (see Fig. 4): North-Central (green), Rocky Mountain (yellow), Coastal-Cascade (red) and Central Oregon Highlands (purple). All individuals demonstrating cyto-nuclear discordance were found along contact zones between lineages.

Highlands groups), although several individuals from populations close to the geographic transition between AFLP groups demonstrated a mismatch between mtDNA group and AFLP cluster membership (Figs 2 and 5). The sole sample from the currently described Santa Cruz

subspecies came out as a separate network, as did two individuals from south western Oregon.

Haplotype network estimation with fewer characters (to permit the inclusion of an individual from the Sierras in California) resulted in similar relationships to the analysis with the larger character set. However, individuals from south western Oregon were distantly connected to the Coastal-Cascade network rather than representing a distinct network (Fig. 4). The haplotype from the Sierras also connected to the Coastal-Cascade network with a large number of mutational steps (Fig. 4). Both iterations of the analysis also suggest divergence between northern and southern haplotypes within networks (Fig. 4).

Degree of divergence

Based on our most complete survey of the genome (751 AFLP loci for 70 individuals), lineages differed by up to 133 AFLP bands (Table 2). Analysis of molecular variance (amova) suggests that a significant proportion of variation in the AFLP data is attributed to differences between lineages (Percent Variation = 14.9, d.f. = 3, SSD = 570.33, $P \ll 0.05$); although most of the variation observed was due to differences within lineages (Percent Variation = 85.0, d.f. = 65, SSD = 3281.11, $P \ll 0.05$). Average uncorrected pairwise sequence divergence between mitochondrial groups ranged from 2.4% to 5.2% [2.5% to 5.8% when corrected according to Tamura & Nei,
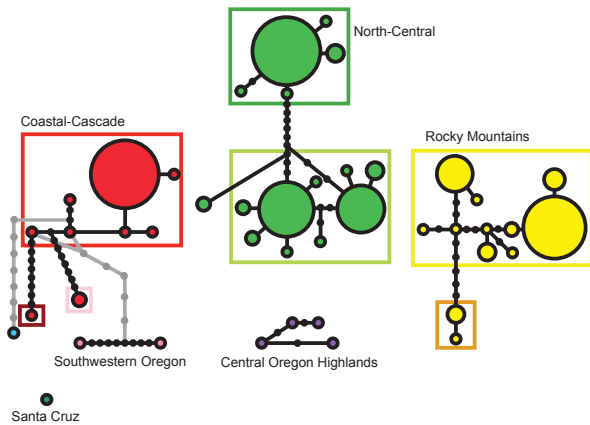
Fig. 4 Haplotype network based on 600 bp of mitochondrial cytochrome b from 142 Long-Toed salamanders. Network estimation was conducted in TCS (Clement *et al.*, 2000) using statistical parsimony. Coloured circles represent haplotypes sampled from different parts of the species' range (see also Fig. 5), with circle diameter proportional to frequency. Missing intermediate haplotypes are shown as small black dots. Groups of haplotypes that are not joined by lines could not be connected within the 95% limits of statistical parsimony. Light grey lines show additional connections that are made when the analysis is repeated with 489 bp to include an individual from the Sierra Mountains in California (teal coloured haplotype). Coloured boxes around groups of haplotypes highlight divergent haplotypes groups within networks and show that further structure is detected within lineages (see also Fig. 5).

(1993); Table 3]. Sequence divergence within groups was substantially lower ranging from 0.06% to 1.8%.

Discussion

Results from our genome-wide survey using AFLPs and from our mitochondrial data set are highly concordant and clearly support the existence of multiple lineages of long-toed salamander. The distributions of these lineages and degree of divergence among them indicate that several refugial populations contributed to the large contemporary range of the species. Furthermore, the relatively restricted geographic areas occupied by the individual long-toed salamander lineages highlight the potential for there to be large discrepancies between estimates of geographic range size depending on the taxonomic treatment of cryptic lineages.

Diversity within the long-toed salamander

Our analysis of several hundred AFLP loci and mtDNA sequence data support at least four major lineages of long-toed salamander. These results thus complement previous studies (Mittleman, 1948; Ferguson, 1961; Thompson, 2003; Thompson & Russell, 2005; Savage, 2008), providing genome-wide evidence of genetic
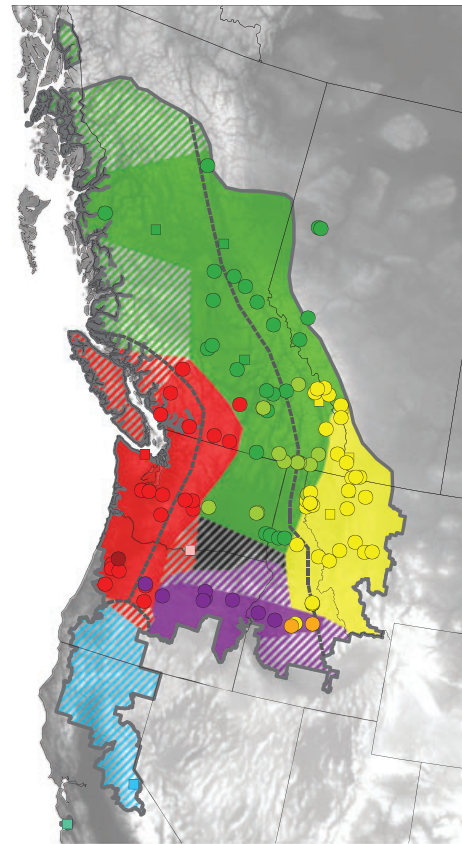

Fig. 5 Approximate distributions of the major lineages of long-toed salamander. Solid coloured areas denote the boundaries of the lineages based on the STRUCTURE analyses of the AFLP data (i.e. Figs 2 and 3) with the lineages coloured as follows: red = Coastal-Cascade, green = North-Central, yellow = Rocky Mountains, purple = Central Oregon Highlands. Striped colours represent regions where AFLP sampling was sparse and the distributions of the lineages are more uncertain and based only on mtDNA, geophysiological features and/or previously described morphological variation (e.g. Ferguson, 1961). Blue striping denotes the described range of the putative *A. m. sigillatum* subspecies and the teal point sample denotes the location of *A. m. croceum*. The black striped area represents an area where the species is very sparsely distributed. Sampling locations are coloured according the mtDNA group to which the majority of individuals from each site belong, with different shades denoting divergent groups within each network as indicated in Fig. 4. Several instances of cyto-nuclear discordance are thus made evident and establish that hybridization occurs where the lineages come into contact.

differences between some of the described subspecies. Our results also contribute much needed clarification of the boundaries of these and other cryptic groups within the species, highlighting several key places where described subspecies' boundaries are inaccurate or do not adequately capture diversity within the species (see Fig. 5 for summary).

Table 2 Average pairwise AFLP distances (number of bands that differ in presence/absence) within (diagonal) and between groups.

	Coastal-Cascade	Central-Oregon Highlands	Rocky Mountains	North-Central
Coastal-Cascade	117.0			
Central-Oregon Highlands	133.3	110.0		
Rocky Mountains	128.3	126.8	102.1	
North-Central	128.0	124.7	106.3	90.3

One of the most notable results of our genetic survey was observation of a cryptic lineage in the southern portion of the currently described range of *A. m. columbianum* (Fig. 5). Individuals from this region come out as a distinct group in STRUCTURE analyses of $K = 5$ (the optimal K) and harbour unique mtDNA haplotypes. These individuals cluster closely with Coastal-Cascade individuals in the MDS analysis (at least in the two dimensions considered presently) and when the STRUCTURE analysis is run at $K = 4$. Intriguingly, however, mtDNA haplotypes from this region appear most closely related to haplotypes from the North-Central portion of the species' range. When the number of mutational steps permitted in the haplotype network analyses is increased to 12, beyond the limits of statistical parsimony, haplotypes from the Central Oregon Highlands form a distantly related group within the North-Central haplotype network. Pairwise sequence divergence between these two groups is also low (Table 3). Previous analyses of mtDNA data have placed two samples from central Oregon into a clade encompassing *A. m. columbianum* (Savage, 2008). Thus, populations from the Central Oregon Highlands have an interesting history, with most analyses distinguishing these populations as a unique lineage but with different markers, suggesting different histories of association with other lineages.

Our mitochondrial data point to additional genetic structure across the species' range. Most notably, the haplotype from the sole individual representing the very spatially restricted Santa Cruz isolate could not be connected to any other network within the limits of statistical parsimony in our haplotype network analysis. High levels of sequence divergence between this haplotype and haplotypes from other places (Table 3) suggest that the Santa Cruz region harbours a distinct group of long-toed salamanders (see also Savage, 2008). Likewise, an individual from the Sierra Nevada Mountains, representing what is currently described as *A. m. sigillatum*, was highly distinct in our haplotype network. However, the degraded quality of DNA from these two individuals precluded us from obtaining AFLP data to test whether these patterns are representative of the entire genome; more extensive sampling of the southernmost portion of the species' range is required before such conclusions can be reached. Such rigorous assessment is particularly pertinent given several examples where genetic structure is apparent in the mitochondrial data but not in the AFLP data. For instance, samples from south western Oregon come out as a distinct haplotype network (Fig. 4) but have AFLP profiles that fall clearly within the Coastal-Cascade lineage. Likewise, none of the distinct groups within each major haplotype network (Fig. 4) were observed in any of the AFLP analyses, suggesting that much of the population structure observed in the mtDNA reflects more recent divergence.

Widespread species as a patchwork of the post-glacial ranges of many populations

Results from a recent study by Shafer *et al.* (2010), demonstrating a positive association between geographic range size and persistence in single vs. multiple glacial refugia, highlight the potential for colonization from multiple refugia to have been important for the generation of large geographic ranges in northern areas. The long-toed salamander was one of the most widespread species featured in that analysis; yet, conclusions about the number of refugial populations were based on a very limited data set (288 bp of mtDNA sequence data) that supported only two lineages of long-toed salamander in previously glaciated areas (see fig. 2 of Thompson & Russell, 2005). Our genome-wide survey

Table 3 Average uncorrected (top) and TN93-corrected (Tamura & Nei, 1993; bottom) pairwise sequence divergence long-toed salamander lineages identified using statistical parsimony. Diagonals show average uncorrected/corrected values within lineages.

	Santa Cruz	South western Oregon	Coastal-Cascade	Central Oregon Highlands	Rocky Mountains	North-Central
Santa Cruz	NA/NA*	0.030	0.034	0.052	0.049	0.043
South western Oregon	0.032	0.018/0.019	0.024	0.043	0.047	0.037
Coastal-Cascade	0.037	0.025	0.009/0.009	0.038	0.042	0.032
Central Oregon Highlands	0.058	0.048	0.042	0.006/0.006	0.049	0.029
Rocky Mountains	0.056	0.056	0.046	0.056	0.009/0.010	0.038
North-Central	0.048	0.048	0.035	0.031	0.041	0.012/0.012

*The putative Santa Cruz subspecies was represented by a single individual in the analysis and is shown in pairwise comparisons for illustration purposes only.

of diversity points to the existence of at least one additional lineage of long-toed salamander in the north (see also Savage, 2008), lending even more support to the association between range size and number of refugia observed by Shafer *et al.* (2010) and highlighting the potential for high-resolution markers to uncover diversity within widespread species that may have been missed by studies using a limited number of markers.

With respect to the history of the species' range, several lines of evidence indicate that the long-toed salamander maintained populations in multiple glacial refugia during the Pleistocene. Informal estimates of divergence using a rate of change of 0.7–1% sequence divergence per million years (based on molecular clocks calibrated using other salamander species as reviewed by Caccone *et al.*, 1997) suggest that the split between the most closely related long-toed salamander lineages occurred at least 2.4 million years ago. Although informal, these estimates preclude divergence after the last glacial maximum (i.e. ~20 000 ybp) and thus rule out a single post-glacial colonization event followed by diversification. Likewise, the average number of AFLP bands (i.e. SNPs) that differed between lineages was high (Table 2), suggesting an extended history of population isolation. Necessarily, all of the lineages described here have clear overlap with many of the previously proposed glacial refugia in the Pacific Northwest (Fig. 1: inset and Table 4). Thus, rather than representing an impressive feat of colonization of the Pacific Northwest by a single lineage, the large contemporary range of the long-toed salamander complex is best explained by the long-term persistence of at least four geographically separated populations during the Pleistocene glaciations and the collective post-glacial expansions of these individual populations. Such fusing of glacial isolates may have been critical in allowing some northern taxa to become widespread, especially species such as the long-toed salamander for which there is evidence of limited dispersal capabilities and a sensitivity to dispersal barriers (Tallmon *et al.*, 2000; Giordano *et al.*, 2007; Goldberg & Waits, 2010; Savage *et al.*, 2010) that might otherwise have hindered colonization of the full range.

Taxonomy and implications for geographic range size

Formal taxonomic evaluation of the long-toed salamander is beyond the scope of this study. However, we note that taxonomic treatment of the different lineages fundamentally affects estimates of geographic range size. As a single species, the described geographic range of the long-toed salamander is approximately 13×10^5 km^2 (based on the range map provided by the Global Amphibian Assessment Database: IUCN, Conservation International & Natureserve, 2008). In contrast, the geographic range sizes of the individual long-toed salamander lineages are magnitudes lower, varying from approximately 135 (putative Santa Cruz lineage supported by the mtDNA) to 58×10^4 km^2 (North-Central). The latter estimates are well within the distribution of range sizes of other amphibians in the Pacific Northwest. Descriptions of the long-toed salamander as a 'widespread species' thus depend critically on taxonomic treatment of the different lineages.

Whether the different long-toed salamander lineages should be considered separate species depends on the species' concept employed (see Coyne & Orr, 2004 for review of species' concepts). Genome-wide differences between the lineages and levels of mtDNA sequence divergence consistent with what has been reported among young amphibian sister species (giant salamanders in the genus *Dicamptodon*, 4.3–6.7%: Steele *et al.*, 2005; torrent frogs in the genus *Amolops*, < 1–3.1%: Matsui *et al.*, 2006) would argue for splitting the species under many genetic definitions [although we note that others have reported similar levels of mtDNA variation between what are still considered subspecies within species (e.g. toads in the *Bufo americanus* complex, 1.8–3.96%: Masta *et al.*, 2002)]. Morphological differences (Ferguson, 1961) and the variation in life history (Anderson, 1967;

Table 4 Overlap between different lineages of long-toed salamander and previously proposed glacial refugia in the Pacific Northwest (see Fig. 1 inset for map).

Lineage	Putative Refugia
Coastal-Cascade	Olympic Peninsula (Soltis *et al.*, 1997)
	Columbia River Drainage (Wagner *et al.*, 2005; Steele & Storfer, 2006)
	Coastal Mountains (Godbout *et al.*, 2008)
	Klamath-Siskiyou Mountains (Soltis *et al.*, 1997; Kuchta & Tan, 2005; Steele & Storfer (2006)
North-Central (Northern and south-eastern)*	Clearwater River Drainage (Carstens *et al.*, 2004)
	Edge of Columbia Plateau (Godbout *et al.*, 2008)
Central Oregon Highlands	Blue-Wallowa Mountains west of the Snake River (see also Thompson & Russell, 2005)
Rocky Mountain	Clearwater River Drainage (Carstens *et al.*, 2004; Nielson *et al.*, 2006)
	Salmon River Drainage (Carstens *et al.*, 2005; Nielson *et al.*, 2006)

*Note that the presence of divergent haplotypes of North-Central salamanders in an isolated population just east of the Cascades in Washington also points to long-term persistence of the lineage in the Columbia Plateau, an area that is largely a gap in the distribution of the species (and in the distributions of other species) apart from a few isolated populations.

see also Kezer & Farner, 1955; Howard & Wallace, 1985) and diet (Anderson, 1968) that have been reported among some subspecies would also support the recognition of distinct species from an ecological standpoint. However, our data do reveal several cases of cytonuclear discordance in the contact zones between lineages, indicating that some hybridization does occur. Our genetic survey suggests that the extent of such cytonuclear discordance and hybridization is generally limited to narrow (50–125 km wide) contact zones between lineages (see Fig. 5). Nevertheless, fine-scale genetic sampling across the lineage boundaries identified presently as well as experimental tests of reproductive isolation between lineages are necessary to further characterize these contact zones and determine whether sufficient reproductive isolation exists to consider these lineages species under the widely employed biological species concept (Coyne & Orr, 2004).

Although the taxonomic status of the different long-toed salamander lineages requires further evaluation, the finding that the most widely distributed salamander in the Pacific Northwest consists of distinct, relatively geographically restricted genetic groups underscores the value of phylogeographic assessment of species' range limits. Several other phylogeographic studies reveal cases where cryptic genetic groups—should they warrant taxonomic recognition—necessitate significant reductions to the described ranges of widespread species (e.g. Irwin *et al.*, 2001; Toews & Irwin, 2008; Oliver *et al.*, 2009; Tan *et al.*, 2010). Such overestimates of species' ranges may be common if 'widespread species' in northern hemispheres generally do have a history of persistence in multiple glacial refugia and thus represent patchworks of potentially very divergent lineages (e.g. Berggren *et al.*, 2005; Niedzialkowska *et al.*, 2011; additional examples in Shafer *et al.*, 2010). Identification and taxonomic scrutiny of genetic lineages within these species is pertinent for interpreting the results of the many studies that use published estimates of species' distributions (e.g. NatureServe) to explore hypotheses concerning geographic range size.

Outstanding questions

That some widespread species represent patchworks of distinct populations that have withstood the test of time raises a number of additional questions. Critically, postglacial colonization from multiple refugia may explain the large contemporary distribution of some northern species such as the long-toed salamander; however, what allowed these species to initially maintain populations in widely separated refugia? Did species with currently restricted ranges that suggest a history of expansion from a single refugium once have wider ranges and simply suffer greater extinction during the Pleistocene glaciations? If so, what factors explain such variation in persistence during environmental change

(e.g. Davies *et al.*, 2009; Waldron, 2010)? Additionally, are widespread species from northern areas more likely to reflect a patchwork of highly divergent genetic groups than widespread species found in regions that were affected by less extreme climatic and habitat change during the Pleistocene (but see Pfenninger & Schwenk, 2007)?

Within widespread taxa that demonstrate marked genetic structure, addressing the factors limiting the spread of individual lineages is also of interest. For instance, in the case of the long-toed salamander, there is considerable variation in the distribution of individual lineages, raising questions as to why some lineages were able to colonize a much bigger area than others (e.g. the North-Central lineage, Fig. 5) and whether these lineages might have been able to eventually occupy the entire contemporary range of the long-toed salamander if the other lineages were not present. Data on dispersal barriers, ecological divergence and hybridization dynamics between the individual lineages are necessary to address these questions.

Finally, we note that not all widespread taxa demonstrate strong phylogeographic structure. In the case of amphibians, some of the most widespread species demonstrate remarkably little genetic differentiation across their range (e.g. *Ambystoma maculatum*: Zamudio & Savage, 2003; *Lithobates [Rana] sylvatica*: Lee-Yaw *et al.*, 2008; *Gastrophryne carolinensis*: Makowsky *et al.*, 2009), including one of the few others to be widely distributed throughout the Pacific Northwest (*Anaxyrus [Bufo] boreas*: Goebel *et al.*, 2009). Continued advances in genomic techniques will not only allow us to better survey taxa and thus identify true outliers with respect to range size but will provide new means for testing ecological and evolutionary hypotheses for the remarkable variation observed in species' distributions.

Acknowledgments

311931-2010 and an Alberta Conservation Association Grant in Biodiversity.

References

Anderson, J.D. 1967. A comparison of the life histories of coastal and montane populations of *Ambystoma macrodactylum* in California. *Am. Midl. Nat.* **77**: 323–355.

Anderson, J.D.. 1968. A comparison of the food habits of *Ambystoma macrodactylum sigillatum, Ambystoma macrodactylum croceum,* and *Amybstoma tigrinum californiense. Herpetologica* **24**: 273–284.

Arnason, U., Gullberg, A., Janke, A., Joss, J. & Elmerot, C. 2004. Mitogenomic analysis of deep gnathostome divergences: a fish is a fish. *Gene* **333**: 61–70.

Arrigo, N., Holderegger, R. & Alvarez, N. 2012. Automated scoring of AFLPs using RawGeno v 2.0, a free R CRAN library. *Methods Mol. Biol.* **888**: 155–175.

Barber, B.R. & Jensen, G. 2012. Quaternary climate change was not an engine of diversification in New World bats (Chiroptera). *J. Mamm. Evol.* **19**: 129–133.

Berggren, K.T., Ellegren, H., Hewitt, G.M. & Seddon, J.M. 2005. Understanding the phylogeographic patterns of European hedgehogs, *Erinaceus concolor* and *E. europaeus* using the MHC. *Heredity* **95**: 84–90.

Brunsfeld, S.J., Sullivan, J., Soltis, D.E. & Soltis, P.S.. 2001. Comparative phylogeography of northwestern North America: a synthesis. In: *Integrating Ecological and Evolutionary Processes in a Spatial Context* (J. Silverton & J. Antonovics eds), pp. 319–339. Blackwell Science, Oxford.

Caccone, A., Milinkovitch, M.C., Sbordoni, V. & Powell, J.R. 1997. Mitochondrial DNA rates and biogeography in European newts (Genus *Euproctus*). *Syst. Biol.* **46**: 126–144.

Cambefort, Y. 1994. Body-size, abundance, and geographical-distribution of Afrotropical dung beetles (Coleoptera, Scarabaeidae). *Acta. Oecol.* **15**: 165–179.

Carstens, B.C., Stevenson, A.L., Degenhardt, J.D. & Sullivan, J. 2004. Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. *Syst. Biol.* **53**: 781–792.

Carstens, B.C., Degenhardt, J.D., Stevenson, A.L. & Sullivan, J. 2005. Accounting for coalescent stochasticity in testing phylogeographical hypotheses: modeling Pleistocene population structure in the Idaho giant salamander *Dicamptodon aterrimus. Mol. Ecol.* **14**: 255–265.

Clement, M., Posada, D. & Crandall, K. 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**: 1657–1660.

Coyne, J.A. & Orr, H.A.. 2004. *Speciation*. Sinauer Associates, Inc., Sunderland.

Davies, T.J., Purvis, A. & Gittleman, J.L. 2009. Quaternary climate change and the geographic ranges of mammals. *Am. Nat.* **174**: 297–307.

Dowling, D.K., Friberg, U. & Lindell, J. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends Ecol. Evol.* **23**: 546–554.

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**: 2611–2620.

Excoffier, L., Laval, G. & Schneider, S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online.* **1**: 47–50.

Falush, D., Stephens, M. & Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* **7**: 574–578.

Ferguson, D.E. 1961. The geographic variation of *Ambystoma macrodactylum* Baird, with the description of two new subspecies. *Am. Midl. Nat.* **65**: 311–338.

Garcia-Barros, E. & Romo Benito, H. 2010. The relationship between geographic range size and life history trait: is biogeographic history uncovered? A test using the Iberian butterflies *Ecography* **33**: 392–401.

Gaston, K.J.. 2003. *The Structure and Dynamics of Geographic Ranges*. Oxford University Press, Oxford.

Gaston, K.J. & Blackburn, T.M. 1996. Conservation implications of geographic range size-body size relationships. *Conserv. Biol.* **2**: 638–646.

Giordano, A.R., Ridenhour, B.J. & Storfer, A. 2007. The influence of altitude and topography on genetic structure in the long-toed salamander (*Ambystoma macrodactylum*). *Mol. Ecol.* **16**: 1625–1637.

Godbout, J., Fazekas, A., Newton, C.H., Yeh, F.C. & Bousquet, J. 2008. Glacial vicarince in the Pacific Northwest: evidence from a lodgepole pine mitochondrial DNA minisatellite for multiple genetically distinct and widely separated refugia. *Mol. Ecol.* **17**: 2463–2475.

Goebel, A.M., Ranker, T.A., Corn, P.S. & Olmstead, R.G. 2009. Mitochondrial DNA evolution in the *Anaxyrus boreas* species group. *Mol. Phylogenet. Evol.* **50**: 209–225.

Goldberg, C.S. & Waits, L.P. 2010. Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Mol. Ecol.* **19**: 3650–3663.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acid. S.* **41**: 95–98.

Hewitt, G. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos. T. R. Soc. Lon. B.* **359**: 183–195.

Howard, J.H. & Wallace, R.L. 1985. Life history characteristics of populations of the Long-toed salamander (*Ambystoma macrodactylum*) from different altitudes. *Am. Midl. Nat.* **113**: 361–373.

Irwin, D.E. 2002. Phylogeographic breaks without geographic barriers to gene flow. *Evolution* **56**: 2383–2394.

Irwin, D.E.. 2012. Local adaptation along smooth ecological gradients causes phylogeographic breaks and phenotypic clustering. *Am. Nat.*, in press.

Irwin, D.E., Alström, P., Olsson, U. & Benowitz-Fredericks, Z. M. 2001. Cryptic species in the genus *Phylloscopus* (Old World leaf warblers). *Ibis* **143**: 233–247.

Irwin, D.E., Rubtsov, A.S. & Panov, E.N. 2009. Mitochondrial introgression and replacement between yellowhammers (*Emberiza citrinella*) and pine buntings (*E. leucocephalos*; Aves, Passeriformes). *Biol. J. Linn. Soc.* **98**: 422–438.

IUCN, Conservation International & Natureserve. 2008. Global amphibian assessment. Available at: http://www.iucnredlist.org/amphibians/

Kezer, J. & Farner, D.S. 1955. Life history patterns of the salamander *Ambystoma macrodactylum* in the high Cascade mountains of southern Oregon. *Copeia* **1955**: 127–131.

Kuchta, S.R. & Tan, A. 2005. Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa. Mol. Ecol.* **14**: 225–244.

Lee-Yaw, J.A., Irwin, J.T. & Green, D.M. 2008. Postglacial range expansion from northern refugia by the wood frog, *Rana sylvatica. Mol. Ecol.* **17**: 867–884.

Leger, E.A. & Forister, M.L. 2009. Colonization, abundance, and geographic range size of gravestone lichens. *Basic Appl. Ecol.* **10**: 279–287.

Lester, S.E., Ruttenberg, B.I., Gaines, S.D. & Kinlan, B.P. 2007. The relationship between dispersal ability and geographic range size. *Ecol. Lett.* **10**: 745–758.

Levsen, N.D., Tiffin, P. & Olson, M.S. 2012. Pleistocene speciation in the genus *Populus* (Salicaceae). *Syst. Biol.* **61**: 401–412.

Loxdale, H.D., Lushai, G. & Harvey, J.A. 2011. The evolutionary improbability of 'generalism' in nature, with special reference to insects. *Biol. J. Linn. Soc.* **103**: 1–18.

Makowsky, R., Chesser, J. & Rissler, L.J. 2009. A striking lack of genetic diversity across the wide-ranging amphibian *Gastrophryne carolinensis* (Anura: Microhylidae). *Genetica* **135**: 169–183.

Masta, S.E., Sullivan, B.K., Lamb, T. & Routman, E.J. 2002. Molecular systematics, hybridization, and phylogeography of the *Bufo americanus* complex in Eastern North America. *Mol. Phylogenet. Evol.* **24**: 302–314.

Matsui, M., Shimada, T., Lui, W., Maryati, M., Khonsue, W. & Orlov, N. 2006. Phylogenetic relationships of Oriental torrent frogs in the genus *Amolops* and its allies (Amphibia, Anura, Ranidae). *Mol. Phylogenet. Evol.* **38**: 659–666.

Mila, B., Carranza, S., Guillaume, O. & Clobert, J. 2010. Marked genetic structuring and extreme dispersal limitation in the Pyrenean brook newt *Calotriton asper* (Amphibia: Salamandridae) revealed by genome-wide AFLP but not mtDNA. *Mol. Ecol.* **19**: 108–120.

Mittleman, M.B. 1948. American Caudata II. Geographic variation in *Ambystoma macrodactylum*. *Herpetologica* **4**: 81–95.

Mueller, R.I., Macey, J.R., Jaekel, M., Wake, D.B. & Boore, J.L. 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. *P. Natl. Acad. Sci.* **101**: 13820–13825.

Niedzialkowska, J., Jedrzejewska, B., Honnen, A., Otto, T., Sideorovich, V.E., Perzanowski, K. *et al.* 2011. Molecular biogeography of red deer *Cervus elaphus* from eastern Europe: insights from mitochondrial DNA sequences. *Acta Theriol.* **56**: 1–12.

Nielson, M., Lohman, K., Daugherty, C.H., Allendorf, F.W., Knudsen, K.L. & Sullivan, J. 2006. Allozyme and mitochondrial DNA variation in the tailed frog (Anura: Ascaphus): The influence of geography and gene flow. *Herpetologica* **62**: 235–258.

Oksanen, J., Blanchet, G.F., Kindt, R., Legendre, P., O'Hara, R.B., Simpson, X.X. *et al.* 2011. Vegan: Community Ecology Package. R package version 1.17-9. http://CRAN.R-project.org/package=vegan

Oliver, P.M., Adams, M., Lee, M.S.Y., Hutchinson, M.N. & Doughty, P. 2009. Cryptic diversity in vertebrates: molecular data double estimates of species diversity in a radiation of Australian lizards (*Diplodactylus*, Gekkota). *P. Roy. Soc. B.* **276**: 2001–2007.

Paradis, E., Claude, J. & Strimmer, K. 2004. APE: analysis of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Pfenninger, M. & Schwenk, K. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evol. Biol.* **7**: 121.

Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Prunier, R. & Holsinger, K.E. 2010. Was it an explosion? Using population genetics to explore the dynamics of a recent radiation within Protea (Proteaceae L.). *Mol. Ecol.* **19**: 3968–3980.

Rauch, E.M. & Bar-Yam, Y. 2004. Theory predicts the uneven distribution of genetic diversity within species. *Nature* **431**: 449–452.

Savage, W.K.. 2008. Landscape genetics and phylogenetic relationships in the *Ambystoma macrodactylum* species complex. Ph.D. Dissertation, The University of California, Davis.

Savage, W.K., Fremier, A.K. & Shaffer, H.B. 2010. Landscape genetics of alpine Sierra Nevada salamanders reveal extreme population subdivision in space and time. *Mol. Ecol.* **19**: 3301–3314.

Shafer, A.B.A., Cullingham, C.I., Cote, S.D. & Coltman, D.W. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Mol. Ecol.* **19**: 4589–4621.

Soltis, D.E., Gizendanner, M.A., Strenge, D.D. & Soltis, P.S. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Syst. Evol.* **206**: 353–373.

Steele, C.A. & Storfer, A. 2006. Coalescent-based hypothesis testing supports multiple Pleistocene refugia in the Pacific Northwest of the Pacific giant salamander (*Dicamptodon tenebrosus*). *Mol. Ecol.* **15**: 2477–2487.

Steele, C.A., Carstens, B.C., Storfer, A. & Sullivan, J. 2005. Testing hypotheses of speciation timing in *Dicamptodon copei* and *Dicamptodon aterriumus* (Caudata: Dicamptodonitdae). *Mol. Phylogenet. Evol.* **36**: 90–100.

Tallmon, D.A., Funk, W.C., Dunlap, W.W. & Allendorf, F.W. 2000. Genetic differentiation among long-toed salamander (*Ambystoma macrodactylum*) populations. *Copeia* **1**: 27–35.

Tamura, K. & Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**: 512–526.

Tan, D.S.H., Ang, Y., Lim, G.S., Ismail, M.R.B. & Meier, R. 2010. From 'cryptic species' to integrative taxonomy: an iterative process involving DNA sequences, morphology, and behaviour leads to the resurrection of *Sepsis pyrrhosoma* (Sepsidae: Diptera). *Zool. Scr.* **39**: 51–61.

Thompson, M.D.. 2003. Phylogeography of the Long-Toed Salamander, *Ambystoma macrodactylum*. M.Sc. Dissertation, The University of Calgary.

Thompson, M.D. & Russell, A.. 2005. Glacial retreat and its influence on migration of mitochondrial genes in the long-toed salamander (*Ambystoma macrodactylum*) in Western North America. In: *Migration of Organisms* (A.M.T. Elewa ed.), pp. 205–246. Springer, Berlin, Heidelberg.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.

Toews, D.P.L. & Irwin, D.E. 2008. Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. *Mol. Ecol.* **17**: 2691–2705.

Vekemans, X., Beauwens, T., Lemaire, M. & Roldan-Ruiz, I. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Mol. Ecol.* **11**: 139–151.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M. *et al.* 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**: 4407–4414.

Voss, S.R. & Shaffer, H.B. 1997. Adaptive evolution via a major gene effect: Paedomorphosis in the Mexican axolotl. *P. Natl. Acad. Sci.* **94**: 14185–14189.

Wagner, R.S., Miller, M.P., Crisafulli, C.M. & Haig, S.M. 2005. Geographic variation, genetic structure, and conservation unit designation in the Larch Mountain salamander (*Plethodon larselli*). *Can. J. Zoolog.* **83**: 396–406.

Waldron, A. 2010. Lineages that cheat death: surviving the squeeze on range size. *Evolution* **64**: 2278–2292.

Weir, J.T. & Schluter, D. 2004. Ice sheets promote speciation in boreal birds. *P. Roy. Soc. B.* **271**: 1881–1887.

Wiens, J.J., Kuczynski, C.A. & Stephens, P.R. 2010. Discordant mitochondrial and nuclear gene phylogenies in emydid turtles: implications for speciation and conservation. *Biol. J. Linn. Soc.* **99**: 445–461.

Zamudio, K.R. & Savage, W.K. 2003. Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution* **57**: 1631–1652.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Summary of locality information for sampled cyt b haplotypes and AFLP data.

Appendix S2 mtDNA RFLP digestion protocol.

Figure S1 Graphical methods used to detect the number of groups (*K*) from STRUCTURE analyses of AFLP loci for the long-toed salamander.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.